Journal of Computational Physics 325 (2016) 116–128

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Importance sampling variance reduction for the Fokker–Planck rarefied gas particle method

B.S. Collyer [a,d,*], C. Connaughton [a,b,d], D.A. Lockerby [c]

[a] Centre for Complexity Science, University of Warwick, Coventry CV4 7AL, UK
[b] Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK
[c] School of Engineering, University of Warwick, Coventry, CV4 7AL, UK
[d] London Mathematical Laboratory, 14 Buckingham Street, London WC2N 6DF, UK

ARTICLE INFO

Article history:
Received 4 December 2015
Received in revised form 4 August 2016
Accepted 6 August 2016

Keywords:
Rarefied gas flows
Fokker–Planck equation
Variance reduction

ABSTRACT

The Fokker–Planck approximation to the Boltzmann equation, solved numerically by stochastic particle schemes, is used to provide estimates for rarefied gas flows. This paper presents a variance reduction technique for a stochastic particle method that is able to greatly reduce the uncertainty of the estimated flow fields when the characteristic speed of the flow is small in comparison to the thermal velocity of the gas. The method relies on importance sampling, requiring minimal changes to the basic stochastic particle scheme. We test the importance sampling scheme on a homogeneous relaxation, planar Couette flow and a lid-driven-cavity flow, and find that our method is able to greatly reduce the noise of estimated quantities. Significantly, we find that as the characteristic speed of the flow decreases, the variance of the noisy estimators becomes independent of the characteristic speed.
© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Recent technological advances have resulted in manufacturing processes that have made possible the production of mechanical devices that operate on the scale of microns and nanometers [1]. Such technologies include lab-on-a-chip devices, micro-heat exchangers, gas chromatographers and micro-jet actuators for control in aerospace. At such small scales, the Navier–Stokes–Fourier (NSF) equations are no longer able to accurately model gas flows, due to the length scales of macroscopic gradients approaching the length of the molecules mean free path, λ. This results in the existence of a region known as the Knudsen layer near solid wall boundaries where the gas is prevented from relaxing to thermodynamic-equilibrium, invalidating the assumption that locally the gas is close to thermal equilibrium required for the NSF equations to be valid.

The Boltzmann equation is a mesoscopic model that is considered to provide the most accurate description of rarefied gases beyond Newton's laws. Before the advent of such small-scale technologies, rarefied gas flows' largest application area was in the field of supersonic atmospheric flows, where the Mach number of flow, $\mathrm{Ma} > 1$. Currently, the prevalent method for numerically approximating the solution to the Boltzmann equation in such regimes is a stochastic particle method called direct simulation Monte Carlo (DSMC) [2,3]. Due to the stochastic nature of the method, DSMC becomes very inefficient for

* Corresponding author at: Centre for Complexity Science, University of Warwick, Coventry CV4 7AL, UK.
E-mail address: benjamin.collyer@gmail.com (B.S. Collyer).

http://dx.doi.org/10.1016/j.jcp.2016.08.008
0021-9991/© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

low-speed flows. Typically the Mach number, $\mathrm{Ma} \ll 1$ for flows within micro and nano technologies, and for a given level of statistical error, the computational costs of DSMC scale as Ma^{-2} [4]. This results in very long computation times for such calculations, and methods that are able to more efficiently solve for low-speed flows are highly desirable.

Currently, there are two methods that are able to greatly reduce the variance of the desired thermodynamic outputs of DSMC calculations. The first, low-variance DMSC (LVDSMC), works by adapting the DSMC collision routine to calculate the evolution of the deviation $f_d = f - f_M$ from a Maxwellian distribution f_M [5]. In low speed flows the deviation from equilibrium is small, allowing for a dramatic decrease in the variance of samples. An alternative method, variance reduced DSMC (VRDSMC), is able to work without significant changes to the DSMC algorithm [6]. The method relies on importance sampling, which allows the algorithm to sample from an equilibrium distribution where the thermodynamic variables are known a priori, to create estimators with smaller variance.

More recently an alternative method to DSMC, where the Boltzmann collision operator is approximated by a Fokker–Planck operator, has been developed and shown to be more efficient than the basic DSMC algorithm [7,8]. Like DSMC, it is solved stochastically using notional particles that represent a certain number of real particles in the gas to be simulated, and as such, the basic algorithm suffers from the same inhibitive scaling with the Mach number. Recently Gorji et al. [9] have proposed a method to reduce the variance of the Fokker–Planck solution algorithm that relies on creating a correlated equilibrium solution using the same set of random numbers that are used in the stochastic solution of the non-equilibrium process. The parallel correlated equilibrium process is in effect a control variate for the non-equilibrium process.

In this paper we develop an importance sampling variance reduction scheme for the Fokker–Planck method and demonstrate its effectiveness in simple test cases. The paper is organised in the following way: in section 2 we introduce the Fokker–Planck model, and the numerical stochastic particle scheme for which we later create new variance reduced estimators. We then outline the general method that allows one to create variance reduced estimators by exploiting known information about how the macroscopic fields behave at equilibrium. In section 3 we describe the variance reduction scheme proposed by Gorji el al. [9], which creates a correlated equilibrium scheme. In section 4 we propose our importance sampling scheme, which we test in section 5 on a homogeneous relaxation, Couette flow and a lid-driven cavity flow. We then compare the importance sampling method against the results obtained by using a correlated equilibrium solution.

2. Background

2.1. The Fokker–Planck collision operator

The Fokker Planck collision operator has appeared in several different contexts, originally derived for the distribution function of a Brownian particle in a fluid [10], but can also be derived from an expansion of a linearised Boltzmann equation, when considering the evolution of density function for a particle in a heat bath [11]. It has been used to model electrons, dense liquids and has received attention for its ability to model rarefied gas flows [12]. Jenny et al. [7] demonstrated very good agreement between the Fokker–Planck model, DSMC and experiment for mean velocities and molecular stresses. More recently it has been extended to describe flows of monatomic gas mixtures [13], diatomic molecules [14] and has been coupled to DSMC [15]. When used as a model for a rarefied gas, it should be considered to be a phenomenological model which retains some of the non-local nature of the collision process described by the Boltzmann equation. The equation for the one-particle distribution function $f(\mathbf{x}, \mathbf{v}, t)$, over a state-space comprised of the position $\mathbf{x} \in \mathbb{R}^3$, velocity $\mathbf{v} \in \mathbb{R}^3$ and time $t \in \mathbb{R}^+$ takes the form:

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f = \mathcal{A}(f) \tag{1}$$

$$:= \frac{1}{\tau} \nabla_{\mathbf{v}} \cdot \left[\mathbf{c} f + RT \nabla_{\mathbf{v}} f \right], \tag{2}$$

where τ is a relaxation time, $\mathbf{c} = \mathbf{v} - \mathbf{u}$ is the local relative molecular velocity, $\mathbf{u}$ is the mean velocity:

$$\mathbf{u}(\mathbf{x}, t) = \frac{1}{\rho} \int \mathbf{v} f(\mathbf{x}, \mathbf{v}, t) \, \mathrm{d}\mathbf{v}, \tag{3}$$

T is the local temperature given by

$$T(\mathbf{x}, t) = \frac{1}{3R\rho} \int c^2 f(\mathbf{x}, \mathbf{v}, t) \, \mathrm{d}\mathbf{v}, \tag{4}$$

and where R is the specific gas constant and ρ is the local density given by

$$\rho(\mathbf{x}, t) = \int f(\mathbf{x}, \mathbf{v}, t) \, \mathrm{d}\mathbf{v}. \tag{5}$$

The collision operator $\mathcal{A}$ has the property of conserving mass, momentum and energy. That is

$$\int \mathcal{A}(f)\psi \, d\mathbf{v} = 0, \tag{6}$$

where $\psi = \{1, \mathbf{v}, v^2\}$ is the set of collisional invariants. The advantage of having a collision operator which can be written as a Fokker–Planck equation is that there exists an equivalent stochastic differential equation (SDE) representation for the dynamics of a random variable $\{\mathbf{X}_t, \mathbf{V}_t\}$ whose distribution f evolves according to (2):

$$\mathrm{d}\mathbf{X}_t = \mathbf{V}_t \mathrm{d}t \tag{7}$$

$$\mathrm{d}\mathbf{V}_t = \frac{1}{\tau}(\mathbf{V}_t - \mathbf{U})\mathrm{d}t + \sqrt{\frac{2RT}{\tau}}\,\mathrm{d}\mathbf{W}_t, \tag{8}$$

where $\mathbf{W}_t$ is a 3-dimensional Wiener process. An efficient scheme for evolving a collection of representative particles with positions and velocities $\{\mathbf{X}^j(t), \mathbf{V}^j(t)\}$, $j = 1 \ldots N$, distributed according to the distribution $f(\mathbf{x}, \mathbf{v}, t)$ in time was devised by *Jenny et al.* [7], and can be summarised as:

$$V_i(t + \Delta t) = V_i(t) - \left(1 - e^{\Delta t/\tau}\right)\left(V_i(t) - U_i(t)\right) + \sqrt{\frac{C^2}{B}}\xi_{1,i} + \sqrt{A - \frac{C^2}{B}}\xi_{2,i} \tag{9}$$

$$X_i(t + \Delta t) = X_i(t) + U_i(t)\Delta t + \tau\left(V_i(t) - U_i(t)\right)\left(1 - e^{-\Delta t/\tau}\right) + \sqrt{B}\xi_{1,i}, \tag{10}$$

where $i = 1, 2, 3$ indexes the dimension,

$$A = RT\left(1 - e^{-2\Delta t/\tau}\right), \tag{11}$$

$$B = RT\tau^2\left(\frac{2\Delta t}{\tau} - \left(1 - e^{-\Delta t/\tau}\right)\left(3 - e^{-\Delta t/\tau}\right)\right), \tag{12}$$

$$C = RT\tau\left(1 - e^{\Delta t/\tau}\right)^2, \tag{13}$$

Δt is the time-step, τ is the relaxation time and ξ are independent standard normal distributed random variables. The spatial domain is discretised into cells, and the expectations of macroscopic quantities of interest are calculated during each time-step for each computational cell. The correct viscosity is obtained by choosing the relaxation time $\tau = 2\mu/p$, where μ is the viscosity and p is the pressure.

2.2. Variance reduction for Monte Carlo sampling

In this section we outline the general framework that allows one to reduce the variance of an estimator of a particular random variable. The basic idea is to exploit information about errors in estimates of known quantities, which can be used in the construction of estimates of unknown quantities. Let us suppose we have a random variable X, and we wish to estimate $\mathbb{E}[X]$, let our estimate of $\mathbb{E}[X]$ be denoted by $\hat{X}$. Let Y be a different random variable with known expectation $\mathbb{E}[Y]$ with an estimator denoted by $\hat{Y}$ ($\hat{Y}$ in this setting is commonly termed a control variate). Then for any $\alpha \in \mathbb{R}$ we can use the identity

$$\mathbb{E}[X] = \mathbb{E}[X + \alpha Y] - \alpha\mathbb{E}[Y], \tag{14}$$

to create a new unbiased estimator for $\mathbb{E}[X]$,

$$X_{VR} = \hat{X} + \alpha\hat{Y} - \alpha\mathbb{E}[Y]. \tag{15}$$

The variance of this estimator is

$$\mathrm{Var}[\hat{X}_{VR}] = \mathrm{Var}[\hat{X}] + \alpha^2\mathrm{Var}[\hat{Y}] + 2\alpha\,\mathrm{Cov}[\hat{X}, \hat{Y}], \tag{16}$$

and if we minimise this over possible choices of α, the minimiser α^* can be calculated and is given by

$$\alpha^* = -\frac{\mathrm{Cov}[\hat{X}, \hat{Y}]}{\mathrm{Var}[\hat{Y}]}. \tag{17}$$

Hence the variance for this choice of α is

$$\mathrm{Var}[\hat{X}_{VR}] = \mathrm{Var}[\hat{X}] - \frac{\mathrm{Cov}[\hat{X}, \hat{Y}]^2}{\mathrm{Var}[\hat{Y}]}. \tag{18}$$

The only condition required for the variance of the estimator to be less than the variance of the original estimator is for $\mathrm{Cov}[\hat{X}, \hat{Y}] > 0$, and so $\hat{X}$ and $\hat{Y}$ being dependent is a necessary condition for variance reduction. This is all supposing that we already know α^* which presupposes that we already know $\mathrm{Cov}[\hat{X}, \hat{Y}]$. In general this is something not known a priori, but can be estimated throughout the simulation. In the next sections we will see how in practice it is possible to exploit this.

3. Parallel process variance reduction

We now briefly describe the method proposed by [9]. The objective of the method is to create a stochastic process $\mathbf{Z}_t$ that is able to run in parallel to the original particle scheme, where crucially, the macroscopic fields are already known. If this is performed in a manner where the parallel process is correlated with the original stochastic process then the variance of the estimators can be reduced in the way described in the previous section. The coupling of the new stochastic process $\mathbf{Z}_t$ to the original SDEs (8) is achieved in the following way:

$$\mathrm{d}\mathbf{X}_t = \mathbf{V}_t \mathrm{d}t \tag{19}$$

$$\mathrm{d}\mathbf{V}_t = \frac{1}{\tau}(\mathbf{V}_t - \mathbf{u})\mathrm{d}t + \sqrt{\frac{2RT}{\tau}}\,\mathrm{d}\mathbf{W}_t \tag{20}$$

$$\mathrm{d}\mathbf{Z}_t = \mathbf{A}\mathrm{d}t + D\,\mathrm{d}\mathbf{W}_t \tag{21}$$

where the coefficients $\mathbf{A}$ and D are chosen to keep the marginal distribution of $(\mathbf{X}_t, \mathbf{Z}_t)$, which we denote $f_0(\mathbf{x}, \mathbf{z}, t)$, as a solution to a Fokker–Planck equation

$$\frac{\partial f_0}{\partial t} + \nabla_{\mathbf{x}} \cdot (\mathbf{U} f_0) = \nabla_{\mathbf{v}} \cdot \left[\mathbf{A} f_0 + \frac{D^2}{2} \nabla_{\mathbf{v}} f_0\right], \tag{22}$$

which when supplemented by appropriate boundary conditions, is solved by a Maxwellian

$$f_M(\mathbf{x}, \mathbf{z}, t) = \frac{\rho(\mathbf{x}, t)}{(2\pi RT_0)^{3/2}} \exp\left(-\frac{z^2}{2RT_0}\right). \tag{23}$$

The way to choose $\mathbf{A}$ and D to ensure that (23) is a solution of (22) is discussed in depth in [9]. The coupling of the processes $\mathbf{Z}_t$ and $\mathbf{V}_t$, by the same Wiener process $\mathbf{W}_t$ requires that when we use stochastic methods to generate numerical solutions, the same set of random numbers is used for both the equilibrium and non-equilibrium processes. This results in the correlation of estimates of expected quantities, allowing the kind of variance reduction outlined in the previous section. In this method, the parallel equilibrium process, with distribution f_M, shares the same density ρ as the non-equilibrium process described which has f as its distribution function, and so the method cannot reduce the noise in the density calculation directly. *Gorji et al.* propose that the density ρ is found from the continuity equation for mass, using conventional finite difference methods. The results they obtain for a homogeneous relaxation, Poiseuille flow and lid-driven cavity flows show the method has the ability to substantially reduce noise. Because this method uses common random numbers to reduce the variance, we will refer to this method as a common random numbers (CRN) method.

4. Importance sampling

The method we propose is an importance sampling scheme. It differs from the importance sampling scheme utilised by VRDMSC [6] in the way that inter-particle collisions are accounted for. The principle that underpins the importance sampling remains the same however. Suppose we are interested in evaluating the expectation of $g(\mathbf{V})$ where $\mathbf{V}$ is a random variable distributed according to the distribution function f, then given N independent samples $\{\mathbf{V}^1, \ldots, \mathbf{V}^N\}$ distributed according to f the following definition gives rise to the estimate:

$$\mathbb{E}_f[g(\mathbf{V})] := \int g(\mathbf{v}) f(\mathbf{v})\, d\mathbf{v} \approx \frac{1}{N} \sum_{i=1}^{N} g(\mathbf{V}^i), \tag{24}$$

which we know from the Central Limit Theorem has an error of order $N^{-1/2}$. We now define a weight function

$$W(\mathbf{v}) := \frac{f_{\mathrm{ref}}(\mathbf{v})}{f(\mathbf{v})}, \tag{25}$$

which is a measure of how likely one is to see a particle with velocity $\mathbf{v}$, relative to how likely one is to observe this particle if it was distributed to a reference density f_{ref}. This definition is well defined if the distribution f is absolutely continuous with respect to f_{ref}, meaning that $f_{\mathrm{ref}}(S) = 0$ whenever $f(S) = 0$ for any subset S of the state-space. This definition can be

viewed as a Radon Nikodym derivative. It can then be observed that the expectation of $g(\mathbf{v})$ with respect to the reference distribution can be estimated using the original samples:

$$\mathbb{E}_{f_{\text{ref}}}[g(\mathbf{V})] = \int f_{\text{ref}}(\mathbf{v})\, g(\mathbf{v})\, d\mathbf{v} \tag{26}$$

$$= \int f(\mathbf{v}) \frac{f_{ref}(\mathbf{v})}{f(\mathbf{v})} g(\mathbf{v})\, d\mathbf{v} \tag{27}$$

$$= \int W(\mathbf{v}) g(\mathbf{v}) f(\mathbf{v})\, dv \tag{28}$$

$$\approx \frac{1}{N} \sum_{i=1}^{N} W(\mathbf{V}^i) g(\mathbf{V}^i). \tag{29}$$

This is significant as it allows one to sample from the reference distribution f_{ref}, using the original set of samples from the distribution f. If the reference distribution is Maxwellian, $f_{\text{ref}} = f_{\text{M}}$ where one knows the thermodynamic fields analytically, then it is possible to create variance reduced estimators as described in section 2. In order to practically apply this, we need a method of evolving the set of weights and velocities $\{\mathbf{V}^i, W^i\}$ in time, where we use the shorthand notation $W^i = W(\mathbf{V}^i)$. For VRDSMC this is possible because it can be shown directly from the Boltzmann equation, that if two particles are accepted to collide with weights W_i and W_j then the post collision weights must be equal to $W^i W^j$ (the weight update rule for rejected particle pairs is also obtained from the Boltzmann equation). Because the Fokker–Planck dynamics have no explicit collisions, a different way to update the weights is needed.

4.1. Weight update rule

Importance weights can be initialised exactly because the initial velocities of the particles are distributed according to a prescribed initial distribution f_0. Hence, we need only consider how weights are updated during each timestep. Let us suppose that a given particle updates its velocity from $\mathbf{V}_t \to \mathbf{V}_{t+\Delta t}$, where $\mathbf{V}_t$ is distributed according to f_t and $\mathbf{V}_{t+\Delta t}$ is distributed according to $f_{t+\Delta t}$, and that we know $W_t = W(\mathbf{V}_t)$. In order to update the weight exactly, one would need to know $f_{t+\Delta t}(\mathbf{V}_{t+\Delta t})$, however this distribution function is unknown. A simple method to estimate the updated weight is to truncate the Taylor expansions of the joint distributions of $\mathbf{V}_t$ and $\mathbf{V}_{t+\Delta t}$

$$f(\mathbf{v}_t, \mathbf{v}_{t+\Delta t}) = f(\mathbf{v}_{t+\Delta t}, \mathbf{v}_{t+\Delta t}) - \Delta t \partial_t f(\mathbf{v}_{t+\Delta t}, \mathbf{v}_{t+\Delta t}) + \ldots + \frac{(-\Delta t)^n}{n!} \partial_{t^{(n)}} f(\mathbf{v}_{t+\Delta t}, \mathbf{v}_{t+\Delta t}) + \ldots \tag{30}$$

$$= f_{t+\Delta t}(\mathbf{v}_{t+\Delta t}) + \mathcal{O}(\Delta t), \tag{31}$$

after their first term (above we have used the fact that the joint distribution of two identical random variables is equal to the marginal). Using this and the form of the joint distribution expressed using the conditional distribution

$$f(\mathbf{v}_t, \mathbf{v}_{t+\Delta t}) = f_{t+\Delta t}(\mathbf{v}_{t+\Delta t}|\mathbf{v}_t) f_t(\mathbf{v}_t), \tag{32}$$

allows the creation of the estimate

$$W_{t+\Delta t} \approx \widehat{W}_{t+\Delta t} := \frac{f_{eq}(\mathbf{V}_{t+\Delta t}|\mathbf{V}_t) f_{eq}(\mathbf{V}_t)}{f_{t+\Delta t}(\mathbf{V}_{t+\Delta t}|\mathbf{V}_t) f_t(\mathbf{V}_t)} \tag{33}$$

$$= \frac{f_{eq}(\mathbf{V}_{t+\Delta t}|\mathbf{V}_t)}{f_{t+\Delta t}(\mathbf{V}_{t+\Delta t}|\mathbf{V}_t)} W_t. \tag{34}$$

This approximation immediately has some desirable properties. Firstly, the error of the approximation decays with Δt. Also, it is possible to calculate this explicitly from the update rule $\mathbf{V}_t \to \mathbf{V}_{t+\Delta t}$ given by equation (9). This conditional distribution will be a Gaussian centred on $\mathbf{V}_t$ plus the deterministic drift, with a temperature dependent variance. Further to this, it has the correct conditional expectation $\mathbb{E}[\widehat{W}_{t+\Delta t}|W_t] = W_t$ when the distribution is stationary. However, the approximation (34) introduces a new source of systematic error into the method, and as a result will bias the estimators. We discuss this further in Section 4.3.

4.2. Weight spreading

The update rule (34) on its own is not a suitable choice, because if such a rule is repeated the variance of this approximation diverges, which is a common problem for this type of particle weight importance sampling method [6,16]. This is a problem, because to reduce the variance of our estimators in a meaningful way, we require the weights to be close to unity. To avoid this problem we use the same kernel density estimator approach as used by the VRDSMC method [6]. This is

a non-parametric alternative to the linear-regression based method proposed by [17]. Kernel density estimation (KDE) is a method that allows one to obtain an estimate $\hat{f}$ of a density function f from samples distributed according to that density function in the following way:

$$\widehat{f}(\mathbf{v}) = \frac{1}{N}\sum_{i=1}^{N} K_r(\mathbf{v}-\mathbf{v}^i), \tag{35}$$

where K_r is a kernel function that integrates over the state-space to 1, and r is a smoothing parameter that controls the width of the kernel function. We use the same spherical kernels as [6]:

$$K_r(\mathbf{v}-\mathbf{v}^i) = \begin{cases} 3/(4\pi r^3) & \text{if } \|\mathbf{v}-\mathbf{v}^i\| < r \\ 0 & \text{otherwise,} \end{cases} \tag{36}$$

which returns the reciprocal of the volume of a sphere of radius r if $\mathbf{v}^i$ lies within the sphere of radius r centred on $\mathbf{v}$, and otherwise returns a zero. If we combine this with (34), the update rule that is obtained is

$$W_{t+1}(\mathbf{V}^i) \approx \frac{\sum_{j=1}^{N} K_r(\mathbf{V_i}-\mathbf{V}^j)\widehat{W}_{t+1}(\mathbf{V}^j)}{\sum_{j=1}^{N} K_r(\mathbf{V_i}-\mathbf{V}^j)} \tag{37}$$

$$= \frac{1}{\left|S_r(\mathbf{V}^i)\right|}\sum_{\mathbf{V}^j\in S_r(\mathbf{V}^i)} \widehat{W}_{t+1}(\mathbf{V}^j), \tag{38}$$

where $S_r(\mathbf{V}^i) = \left\{\mathbf{V}^j : \|\mathbf{V}^j-\mathbf{V}^i\| < r\right\}$ is the set of samples whose members lie within the sphere of radius of r centred on $\mathbf{V}^i$. This KDE step prevents the weights from diverging, making the scheme more stable. Increasing the smoothing parameter r results in an estimator with a smaller variance, however it also increases the bias of the estimation, so ideally r should be chosen to be as small as possible whilst maintaining an acceptable level of variation.

4.3. Bias

The bias introduced by the approximation of the weights given in equation (34) is a source of error not present in the original algorithm. For a fixed interval of time, this is an error that does not decay with the size of the time discretisation. However there is no error in the approximation at equilibrium, so for initial conditions close to equilibrium this bias of the estimators is small. To demonstrate this numerically in 1D, we have used a generalised Gaussian distribution f_g defined by:

$$f_g(v) = \frac{\beta}{2\gamma\Gamma(1/\beta)} e^{-\left(\frac{|v-\mu|}{\gamma}\right)^{\beta}}, \tag{39}$$

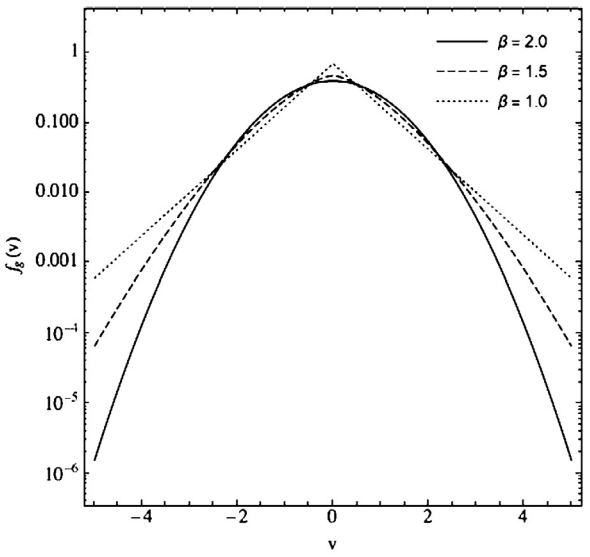
where μ, γ and β are the location, scale and shape parameters, to produce initial distributions of particles with the same mean and variance but differing sizes of kurtosis. For $\beta = 2.0$ the generalised Gaussian distribution reduces to a Gaussian distribution, and for $\beta < 2.0$ produces distributions with heavier tails (see Fig. 1(a)). When left to relax to equilibrium, the kurtosis of these particle distributions will decay to zero, which allows the measurement of the size of the bias in the sample kurtosis, κ, for particle ensembles with initial conditions that vary in their departure from equilibrium.

Fig. 1(b) shows the level of bias measured after the gas is allowed to relax to equilibrium. The estimates at equilibrium were produced after letting the distributions relax for a time period 10τ, with the sample kurtosis averaged over a further 1000 time-steps, and 10 ensembles. The initial conditions were varied by changing the shape parameter β, and adjusting the scale parameter γ so the temperature is fixed. As β approaches 2, the level of bias decreases to zero. The number of time-steps Δ_s between using KDE to stabilise the weights is also varied. The more often the stabilisation is performed, the smaller the bias in the estimator.

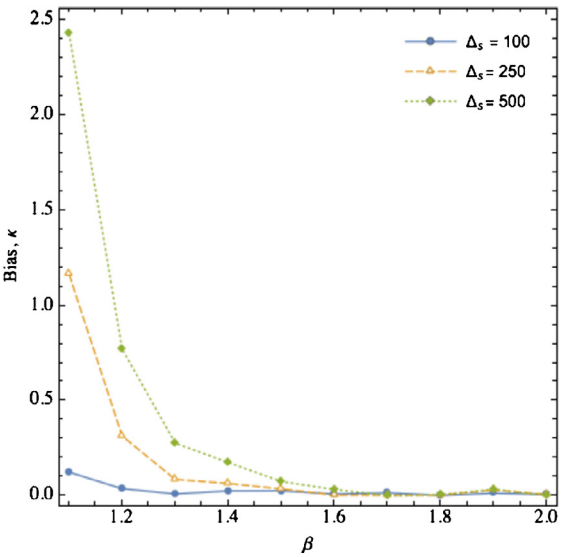
4.4. Boundary conditions

We use the same boundary condition methodology as prescribed by the VRDSMC method, that is for diffusely reflecting fully accommodating walls, with temperature T_{wall} and tangential velocity u_{wall}. Supposing that the Maxwellian distribution at the boundary is given by $f_{wall}(\mathbf{v}) = \rho_{wall}P_{MB}(\mathbf{v})$, where P_{MB} is a Gaussian probability density, and the boundary is the plane $x = 0$, then the no flux boundary condition is given by

$$\rho_{wall}\int_{v_x>0} v_x P_{MB}(\mathbf{v})\,d\mathbf{v} + \int_{v_x<0} v_x f(\mathbf{v})\,d\mathbf{v} = 0, \tag{40}$$



(a) Generalised Gaussian distributions.



(b) Bias in the estimation of the kurtosis at equilibrium.

Fig. 1. The size of bias at equilibrium when measuring the kurtosis κ, for different generalised Gaussian initial distributions. The initial distributions were produced with mean $\mu = 0$ and variance $\sigma^2 = 1$. Simulation parameters: $\tau = 1$, $\Delta_t = 1/100$, relaxation time: 1000 time-steps, averages taken over 1000 time-steps and 10 ensembles.

and similarly for the equilibrium solution

$$\rho_{wall,eq} \int_{v_x>0} v_x P_{MB,eq}(\mathbf{v})\, d\mathbf{v} + \int_{v_x<0} v_x W(\mathbf{v}) f(\mathbf{v})\, d\mathbf{v} = 0. \tag{41}$$

The second integrals in the above equations are the particle fluxes, and can be estimated by counting the number of particles N_{in} that cross through a wall of area Δs in a time period Δt by $(1/\Delta s \Delta t) N_{in}$, and at equilibrium are estimated by $(1/\Delta s \Delta t) \sum_i^{N_{in}} W_i$. Also, we can use the analytical properties of the Gaussian distribution to evaluate the first integrals:

$$\int_{v_x<0} v_x P_{MB}(\mathbf{v})\, d\mathbf{v} = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{kT}{m}}. \tag{42}$$

Therefore, a particle that changes velocity from $\mathbf{V}$ to $\mathbf{V}'$ when colliding with a wall, changes its weight according to

$$W' = W(\mathbf{V}') = \frac{f_{eq}(\mathbf{V}')}{f(\mathbf{V}')} \tag{43}$$

$$= \frac{\rho_{wall,eq} P_{MB,eq}(\mathbf{V}')}{\rho_{wall} P_{MB}(\mathbf{V}')} \tag{44}$$

$$= \sqrt{\frac{T_{wall}}{T_{wall,eq}}} \frac{\sum_i^{N_{in}} W_i}{N_{in}} \frac{P_{MB,eq}(\mathbf{V}')}{P_{MB}(\mathbf{V}')}, \tag{45}$$

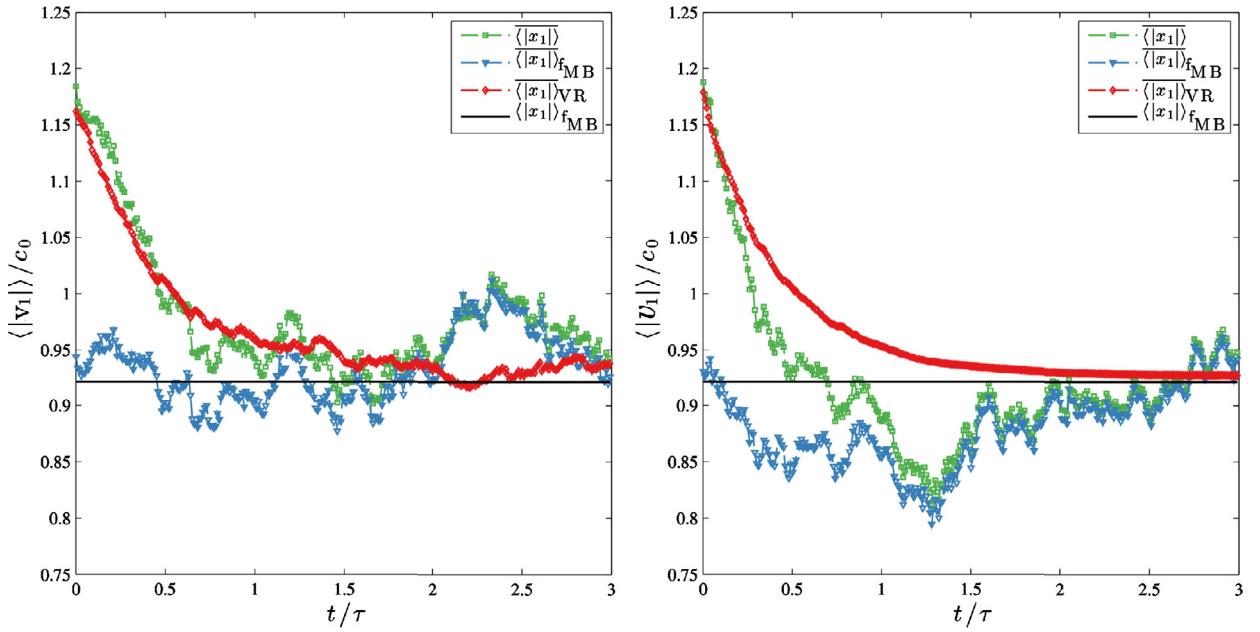
where typically, we choose the temperature of the equilibrium wall boundary condition to be equal to the temperature of non-equilibrium wall boundary condition, i.e. $T_{wall} = T_{wall,eq}$.

5. Results

5.1. Homogeneous relaxation to equilibrium

We will demonstrate the effectiveness of this method first with a homogeneous relaxation to equilibrium, i.e. when $f(t, \mathbf{x}, \mathbf{v}) = f(t, \mathbf{v})$ has no spatial component. We start from an initial distribution of particles

$$f_0(\mathbf{v}) = (1/2)\big(f_M(v_1; c_0, c_0) + f_M(v_1; -c_0, c_0)\big) f_M(v_2; 0, c_0) f_M(v_3; 0, c_0), \tag{46}$$



(a) Without stabilisation (b) With stabilisation

Fig. 2. Homogeneous relaxation towards equilibrium, (a) without KDE, (b) with KDE, smoothing parameter $r = 0.05c_0$. The green squares represent the time-series of the standard MC estimator of the non-equilibrium process; the blue triangles represent the time-series of the standard MC estimator of the process biased to sample from equilibrium; the red diamonds represent the time-series of the VRFP estimator of non-equilibrium process; the black line represents the exact expectation at equilibrium. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

which we know to relax towards the Maxwellian distribution $f_M(\mathbf{v}, \mathbf{0}, \sqrt{(4/3)c_0^2})$. In Figs. 2(a)–2(b) we show how the variance reduced estimator performs against the standard estimator, when estimating $\langle|v_1|\rangle$ using 100 particles, with and without the KDE stabilisation procedure. In both cases, the variance of the new estimator is smaller than the standard estimator, but the estimator with stabilisation from the KDE reduces the variance even further.

5.2. Couette flow

To test the particle weight variance reduction, we have applied the scheme to sample from a steady-state planar Couette flow, and compare to results obtained using a common random number scheme. A Couette flow is a flow where the fluid is bounded by two parallel walls moving in opposite directions within their planes, with velocity $\pm U_{\mathrm{wall}}$. For Knudsen numbers $Kn = 0.05, 0.5, 1.0$ respectively, Figs. 3, 4, 5 show the variance reduced and standard Monte Carlo estimators of the steady-state flow velocity field parallel to the wall, $v_2(x_1)$, (left) as well as the temperature profile across the channel $T(x_1)$, for a Couette flow with wall velocity $v_{\mathrm{wall}} = 0.01c_0$, $\mathrm{Kn} = 0.5$, 20 cells and 100 particles per cell. All the results show a significant improvement of performance over the unweighted standard Monte Carlo estimator.

Next we compare the VRFP importance sampling scheme to the CRN correlated equilibrium scheme. Because we are interested in the noise of the estimate of the velocity profile across the channel, and the speed of the flow is small, we make the simplifying assumption that the steady-state density across the channel is constant. This allows us to choose the coefficients $\mathbf{A} = \mathbf{z}/\tau$ and $D = \sqrt{2RT_{wall}/\tau}$ so that the correlated equilibrium process is distributed according to the global Maxwellian

$$f_M(\mathbf{x}, \mathbf{z}, t) = \frac{\rho}{(2\pi RT_{wall})^{3/2}} \exp\left(\frac{-z^2}{2RT_{wall}}\right). \tag{47}$$

Fig. 6 compares the noise-to-signal ratio of the CRN scheme, VRFP scheme and standard Monte Carlo estimator against signal strength for the samples taken from steady state Couette flow estimator. The results were obtained with parameters $\mathrm{Kn} = 0.5$, 20 cells and 50 particles per cell, with steady-states velocity profiles averaged over 50 time-steps. The results show that the CRN attains a noise-to-signal ratio that is constant over the range of Mach numbers tested, and because of this the relative accuracy over the standard Monte Carlo estimator increases as the Mach number decreases. Similarly, the importance weighted variance reduced estimator has a noise-to-signal ratio that is independent of the signal size as the signal size decreases. The importance sampling estimator, however, achieves a standard deviation which is on average

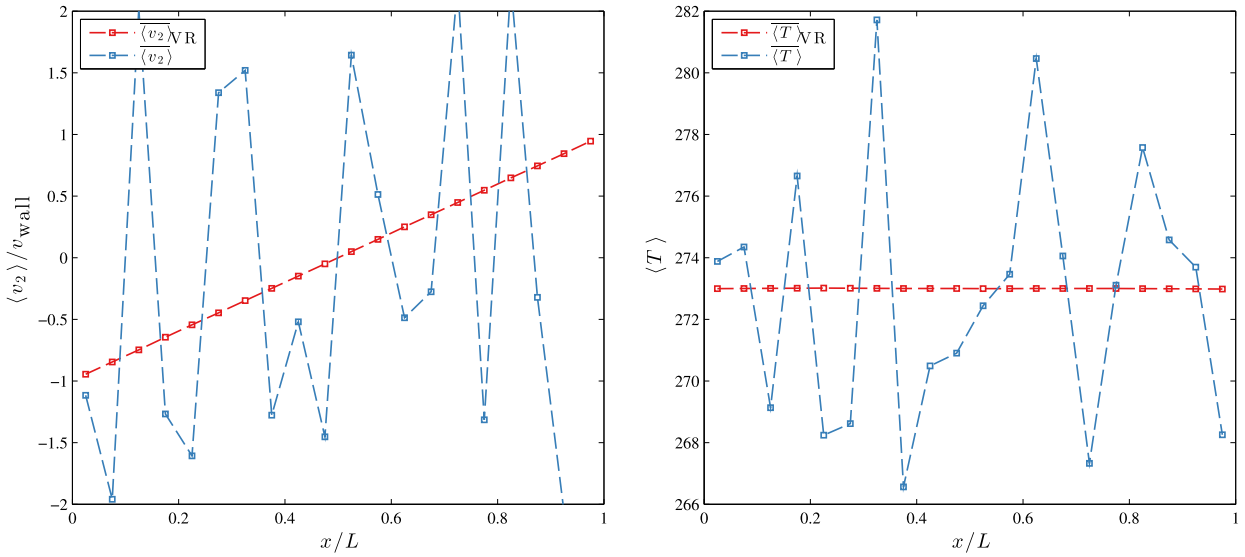


Fig. 3. Couette flow with wall velocity $v_{\text{wall}} = 0.01c_0$, Kn = 0.05, 20 cells and 100 particles per cell.

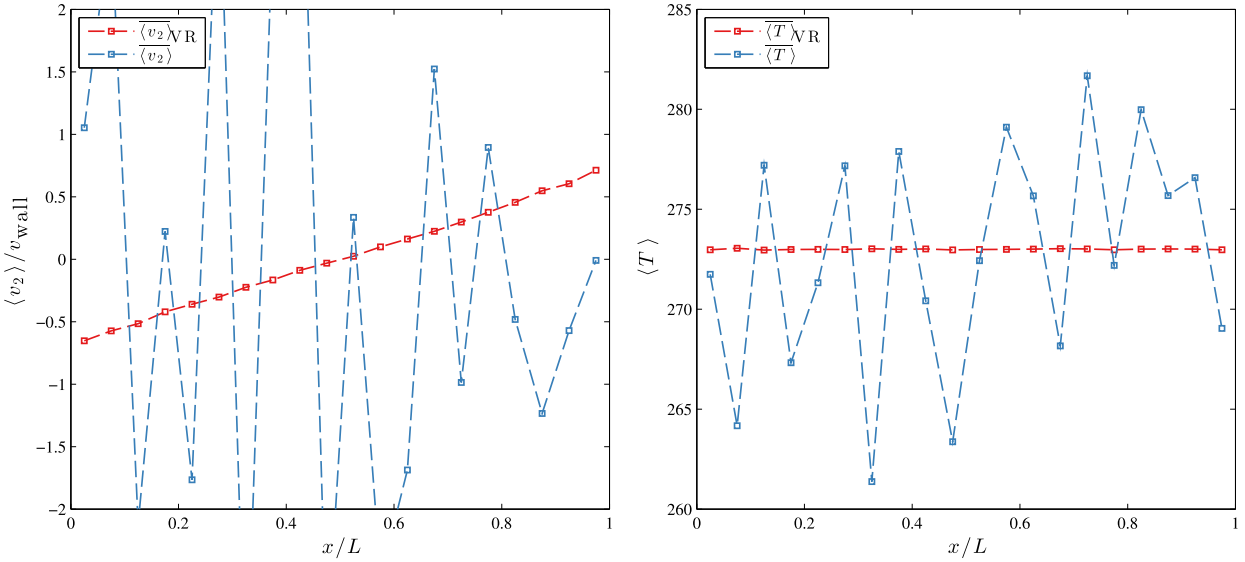


Fig. 4. Couette flow with wall velocity $v_{\text{wall}} = 0.01c_0$, Kn = 0.5, 20 cells and 100 particles per cell.

a factor of 0.24 smaller than the average size of the standard deviation of the CRN estimator. This increase in performance over the CRN scheme is achieved with a time-step with a duration of 0.21 seconds, whereas the CRN scheme had a time-step lasting 0.12 seconds. Hence the VRFP scheme is almost twice as computationally expensive as the CRN scheme.

5.3. Lid-driven cavity

To further demonstrate the effectiveness of the method, we apply it to a lid-driven cavity flow, where the fluid is bounded in two dimensions by a square box in the (x, y) plane, with translational symmetry in the z axis. Three of the bounding walls are stationary, and one of the bounding walls moves within its plane at constant velocity U_{wall}, giving rise to a circulatory flow within the cavity.

Figs. 7(a)–7(b) show the velocity and non-dimensional temperature field $(T/T_0 - 1)$ of the steady state flow, with a lid velocity of $U_{\text{wall}} = 0.001c_0$ for the standard Monte Carlo and variance reduced sampling schemes. The results have been averaged over 5000 time-steps, and 10 independent ensembles on a 50×50 grid, with an average of 30 particles per cell.

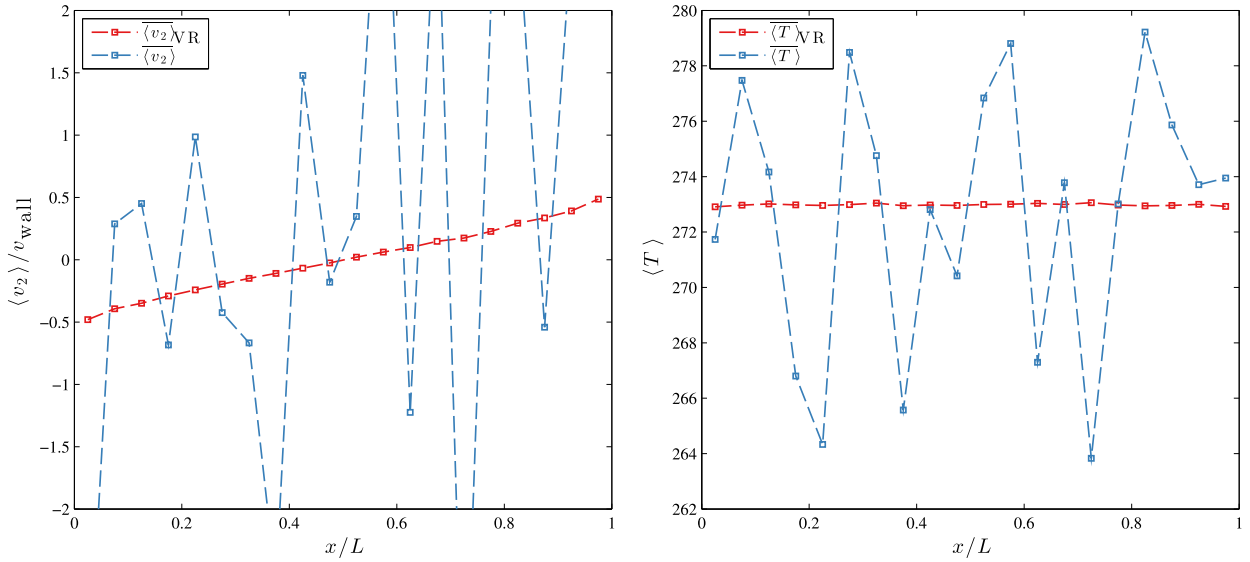


Fig. 5. Couette flow with wall velocity $v_{\text{wall}} = 0.01c_0$, Kn = 1.0, 20 cells and 100 particles per cell.

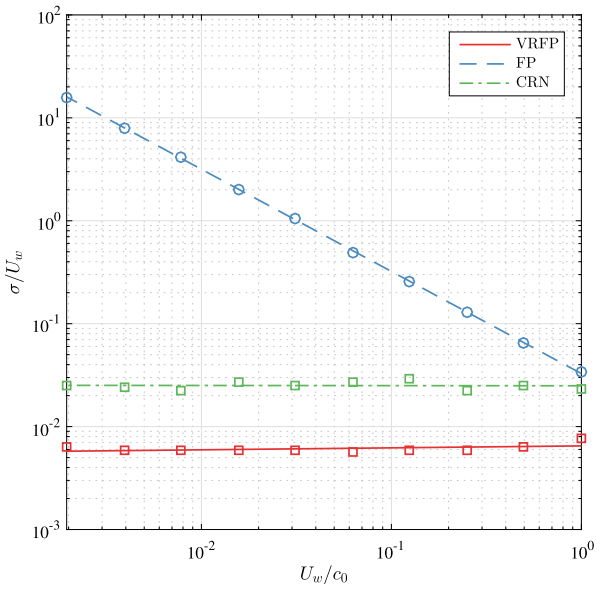


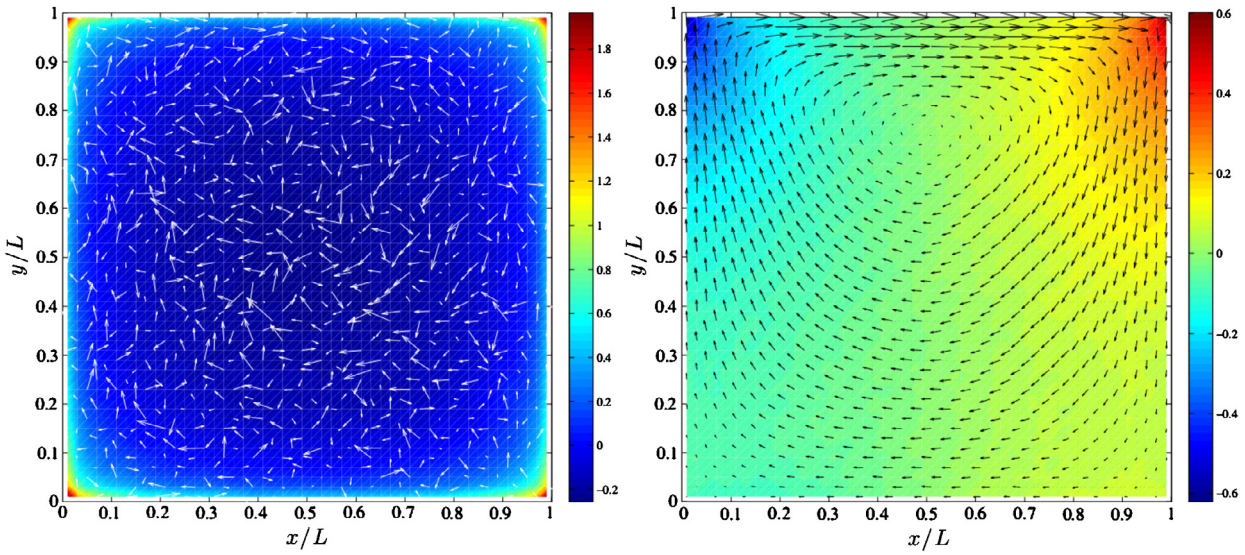
Fig. 6. Comparison of noise-to-signal ratio vs signal size, between standard Monte Carlo, CRN, and our importance sampling method (VRFP).

The standard Monte Carlo scheme is not able to pick up the signal, whereas we see clearly that the importance sampling scheme is able to recover the signal. In Fig. 8 we compute the streamlines of the variance reduced flow, and the shear stress π_{12}.

Fig. 9 shows results from lid-driven cavity flows with lid speeds $0.1c_0$, $0.01c_0$, $0.001c_0$ and $0.0001c_0$. As was the case with the Couette flow, the level of noise in each calculation is independent of the lid-speed.

6. Conclusion

In this paper we have developed an importance sampling method for the Fokker–Planck rarefied gas model, that assigns weights to each stochastic particle allowing one to sample from an equilibrium distribution. We have demonstrated its effectiveness in reducing the variance of estimates of thermodynamics quantities for low Mach number flows over a range of Knudsen numbers. The level of noise in the estimators becomes independent of the Mach number for low-speed flows. Our



(a) Standard Monte Carlo estimate (b) Variance reduced estimate

Fig. 7. Lid-driven cavity flow. Velocity field and non-dimensional temperature $(T/T_0 - 1)/\mathrm{Ma}$. $\mathrm{Kn} = 1.0$, $U_{\mathrm{wall}} = 0.001c_0$, with and without importance sampling variance reduction.

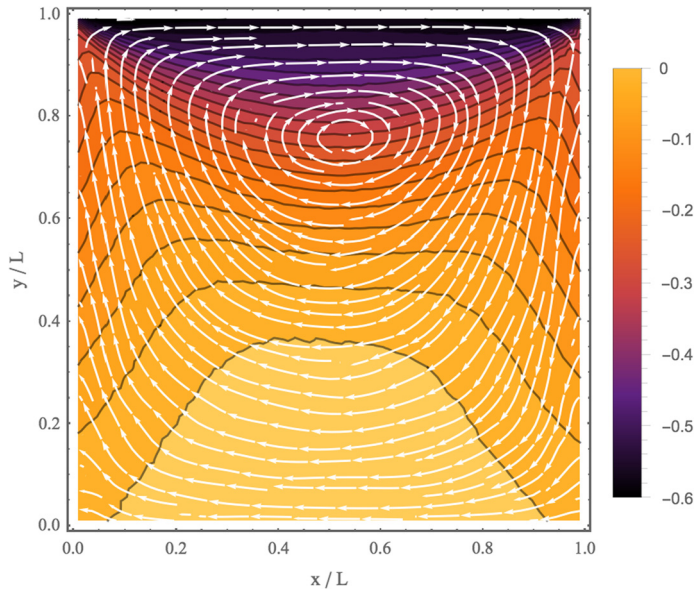
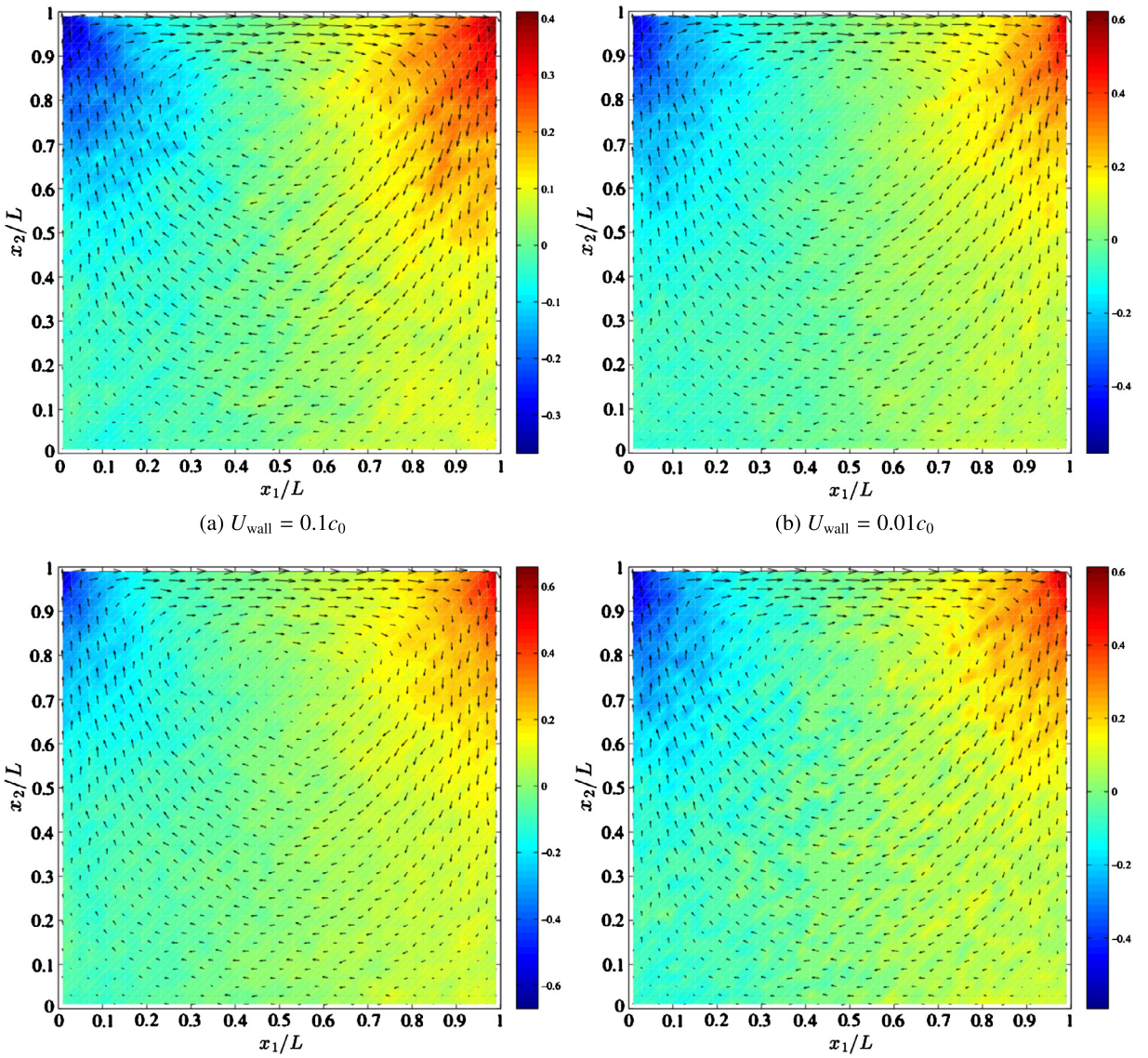


Fig. 8. Lid-driven cavity flow. Streamlines and non-dimensional shear stress $\pi_{12}/(\rho_0 R T_0 \mathrm{Ma})$ of the variance reduced estimate.

results show that although the importance sampling algorithm is more computationally expensive than the CRN method, the relative noise reduction is enough to outweigh the cost. We believe it to be a versatile and robust method, and because it doesn't alter the basic algorithm of the particle solution scheme, can be used in conjunction with other variance reduction schemes such as the CRN method.

Acknowledgements

We thank Dr. Hossein Gorji and coworkers for sending us their paper [9] prior to its finalised publication. This research is financially supported by EPSRC Programme Grants EP/I011927/1 and EP/N016602/1.



(a) $U_{\text{wall}} = 0.1c_0$ (b) $U_{\text{wall}} = 0.01c_0$

(c) $U_{\text{wall}} = 0.001c_0$ (d) $U_{\text{wall}} = 0.0001c_0$

Fig. 9. Lid-driven cavity flows with different wall-speeds. 50×50 grid, 25 particles per cell on average, 5000 time steps to reach steady-state, thermodynamic fields averaged from 5000 further time steps. The level of noise is independent to the wall-speed.

References

[1] J.M. Reese, M.A. Gallis, D.A. Lockerby, New directions in fluid dynamics: non-equilibrium aerodynamic and microsystem flows, Philos. Trans. R. Soc. Lond. A, Math. Phys. Eng. Sci. 361 (1813) (2003) 2967–2988.
[2] G.A. Bird, Molecular Gas Dynamics and the Direct Simulation of Gas Flows, Clarendon Press, 1994.
[3] E.S. Oran, C.K. Oh, B.Z. Cybyk, Direct Simulation Monte Carlo: Recent Advances and Applications, Annu. Rev. Fluid Mech. 30 (1998) 403–441.
[4] N.G. Hadjiconstantinou, A.L. Garcia, M.Z. Bazant, G. He, Statistical error in particle simulations of hydrodynamic phenomena, J. Comput. Phys. 187 (2003) 274–297, arXiv:cond-mat/0207430.
[5] T.M.M. Homolle, N.G. Hadjiconstantinou, A low-variance deviational simulation Monte Carlo for the Boltzmann equation, J. Comput. Phys. 226 (2) (2007) 2341–2358.
[6] H.A. Al-Mohssen, N.G. Hadjiconstantinou, Low-variance direct Monte Carlo simulations using importance weights, ESAIM: Math. Model. Numer. Anal. 44 (5) (2010) 1069–1083.
[7] P. Jenny, M. Torrilhon, S. Heinz, A solution algorithm for the fluid dynamic equations based on a stochastic model for molecular motion, J. Comput. Phys. 229 (4) (2010) 1077–1098.
[8] M.H. Gorji, M. Torrilhon, P. Jenny, Fokker–Planck model for computational studies of monatomic rarefied gas flows, J. Fluid Mech. 680 (2011) 574–601.
[9] M.H. Gorji, N. Andric, P. Jenny, Variance reduction for Fokker–Planck based particle Monte Carlo schemes, J. Comput. Phys. 295 (2015) 644–664.
[10] S. Chandrasekhar, Stochastic problems in physics and astronomy, Rev. Mod. Phys. 15 (1) (1943) 1–89.

[11] L. Ferrari, On the velocity relaxation of a Rayleigh gas: I. Assumptions and approximations in the derivation of the usual kinetic equation, Physica A 115 (1–2) (1982) 232–246.
[12] C. Cercignani, The Boltzmann Equation and Its Applications, Springer, 1988.
[13] H.G. Jenny Patrick, A kinetic model for gas mixtures based on a Fokker–Planck equation, J. Phys. Conf. Ser. 362 (1) (2012) 12042.
[14] M.H. Gorji, P. Jenny, A Fokker–Planck based kinetic model for diatomic rarefied gas flows, Phys. Fluids 25 (6) (2013) 062002.
[15] M.H. Gorji, P. Jenny, An efficient particle Fokker–Planck algorithm for rarefied gas flows, J. Comput. Phys. 262 (2014) 325–343.
[16] J. Chun, D.L. Koch, A direct simulation Monte Carlo method for rarefied gas flows in the limit of small Mach number, Phys. Fluids 17 (10) (2005).
[17] S. Brunner, E. Valeo, J.A. Krommes, Collisional delta-f scheme with evolving background for transport time scale simulations, Phys. Plasmas 6 (12) (1999) 4504–4521.